1

METHOD AND APPARATUS FOR SPEECH ENCODING, SPEECH DECODING, AND SPEECH POST PROCESSING

This application is a division of application Ser. No. 08/243,181, filed May 16, 1994, entitled A METHOD AND APPARATUS FOR SPEECH ENCODING, SPEECH DECODING, AND SPEECH POST PROCESSING and now pending.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a method and apparatus for speech encoding, speech decoding, speech post processing, which are used when speech is transmitted digitally, stored and synthesized.

2. Description of the Related Art

In a conventional speech coding apparatus, input speech taken within analysis windows are analyzed by taking their frequency spectrum. The analysis windows are either aligned with the analysis frames or at a fixed offset from the analysis frames. The analysis frames are defined as having a fixed length and are offset at fixed interval. In a conventional speech decoding apparatus and a speech post processor, the quantization noise of synthesized speech is perceptually reduced by emphasizing peaks (formant) and suppressing other part of the speech spectrum. The peak is produced by the resonation of the vocal tract in the speech spectrum.

An article on the conventional speech coding/decoding apparatus is "Sine-Wave Amplitude Coding at Low Data Rates", (Advance in Speech Coding, Kluwer Academic Publishers, P203–213) of the article 1 by R. Macaulay, T. Parks, T. Quatieri, M Sabin. This article is hereinafter called "article 1". FIG. 12 shows a configuration of the speech coding/decoding apparatus stated in the article 1. The conventional speech coding/decoding apparatus comprises a speech coding apparatus 1, a speech decoding apparatus 2 and a transmission line 3. Input speech 4 is input into the speech coding apparatus 1. Output speech 5 is output from the speech decoding apparatus 2. A speech analysis means 6, a pitch coding means 7, a harmonics coding means 8 are implemented in the speech coding apparatus 1. A pitch decoding means 9, a harmonics decoding means 10, an amplitude emphasizing means 11 and a speech synthesis means 12 are implemented in the speech decoding apparatus 2. The speech coding apparatus 1 has lines 101, 102, 103. The speech decoding apparatus 2 has lines 104, 105, 106, 107.

FIG. 13 shows speech waveforms resulting from operation of the conventional speech coding and decoding apparatus.

The operation of the conventional speech coding/decoding apparatus is described with reference to FIGS. 12 and 13. The input speech 4 is input into the speech analysis means 6 through the line 101. The speech analysis means 6 analyzes the input speech 4 per analysis frame having a fixed length. The speech analysis means 6 analyzes the input speech 4 within an analysis window. The analysis window, that is, for instance, a Hamming window, has its center at the specific location in the analysis frame. The speech analysis means 6 extracts a power P of the input speech within the analysis window. The speech analysis means 6 also extracts a pitch frequency by using, for instance, an auto correlation analysis. The speech analysis means 6 also extracts an

2

amplitude Am and a phase θm (m is a harmonic number) of a harmonic components on a frequency spectrum at an interval of the pitch frequency by a frequency spectrum analysis. FIG. 13 show an example of calculating the amplitude Am of the harmonic components on the frequency spectrum by picking up input speech within one frame. The pitch frequency (1/T, T stands for the pitch length) extracted by the speech analysis means 6 is output to a pitch coding means 7 through the line 103. The power P, and the amplitude Am and the phase θm of the harmonics are output to a harmonics coding means 8 through the line 102.

The pitch coding means 7 encodes the pitch frequency (1/T) input through the line 103 after quantizing. The quantizing is, for example, done using a scalar quantization. The pitch coding means 7 outputs a coded data to the speech decoding apparatus 2 through a transmission line 3.

The harmonics coding means 8 calculates a quantized power P' by quantizing the power P input through the line 102. The quantizing is done, for example, using the scalar quantization. The harmonics coding means 8 normalizes the amplitude Am of the harmonic component input through the line 102 by using the quantization power P' to get a normalized amplitude ANm. The harmonics coding means 8 quantizes the normalized amplitude ANm to get a quantized amplitude ANm'. The harmonics coding means 8 quantizes, for example using the scalar quantization, the phase θm input through the line 102 to get a quantized phase θm'. Then the harmonics coding means 8 encodes the quantized amplitude and the quantized phase θm' and outputs the coded data to the speech decoding apparatus 2 through the transmission line 3.

The operation of the speech decoding apparatus 2 explained. The pitch decoding means 9 decodes the pitch frequency of the coded data of the pitch frequency input through the transmission line 3. The pitch decoding means 9 outputs the decoded pitch frequency to a speech synthesis means 12 in the speech decoding apparatus 2 through the line 104.

A harmonics decoding means 10 decodes the power P', and the amplitude ANm' and the phase θm' of the harmonic components, within the coded data input through the transmission line 3 from the harmonics coding means 8. The harmonics decoding means 10 calculates a decoded amplitude Am' by multlplying the amplitude ANm' by P'. The harmonics decoding means 10 outputs these decoded amplitude Am' and phase θm' to an amplitude emphasizing means 11 through the line 105.

The decoded amplitude Am' contains the quantization noise generated by quantizing. Generally, the human ear has a characteristic of perceiving less quantization noise at peaks (formant part) of the frequency spectrum than at bottoms. By using this characteristic, the amplitude emphasizing means 11 reduces the quantization noise to the human ear. As shown in FIG. 14, the amplitude emphasizing means 11 emphasizes the peaks of the decoded amplitude Am' and suppresses other part of Am'. Thus, the amplitude emphasizing means 11 reduces the quantization noise to the human ear. The emphasized amplitude AEm' and the phase θm' are output to a speech synthesis means 12 through the line 106.

Depending upon the input pitch frequency, the emphasized amplitude AEm' of the harmonic components and the phase θm' , the speech syntheses means 12 synthesizes a decoded speech S(t) using the following formula (1). The decoded speech S(t) is output as an output speech 5 through the line 107.